AL/HR-TP-1995-0009

ARMSTRONG LABORATORY

CORRELATION OF GENERAL COGNITIVE ABILITY AND PSYCHOMOTOR TRACKING TESTS

Malcolm James Ree
Thomas R. Carretta

HUMAN RESOURCES DIRECTORATE
MANPOWER AND PERSONNEL RESEARCH DIVISION
7909 Lindbergh Drive
Brooks AFB, Texas 78235-5352

August 1995

Final Technical Paper for Period October 1994-March 1995

AIR FORCE MATERIEL COMMAND
BROOKS AIR FORCE BASE, TEXAS

NOTICES

Publication of this paper does not constitute approval or disapproval of the ideas or findings. It is published in the interest of scientific and technical information (STINFO) exchange.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

Malcolm James Ree

MALCOLM JAMES REE,
Project Scientist

PATRICK C. KYLLONEN
Technical Director
Manpower & Personnel Research Div

Gary D. Zank

GARY D. ZANK, Colonel, USAF
Chief
Manpower and Personnel Research Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1995	Interim Paper - October 1994 - March 1995

4. TITLE AND SUBTITLE

Correlation of General Cognitive Ability and Psychomotor Tracking Tests

5. FUNDING NUMBERS

PE - 62205F
PR - 7719
TA - 25
WU - 01

6. AUTHOR(S)

Malcolm James Ree
Thomas R. Carretta

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Armstrong Laboratory
Human Resources Directorate
Manpower and Personnel Research Division
7909 Lindbergh Drive
Brooks AFB, Texas 78235-5352

8. PERFORMING ORGANIZATION REPORT NUMBER

AL/HR-TP-1995-0009

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Previously published in the International Journal of Selection and Assessment, 1994.
Armstrong Laboratory Technical Monitor: Dr Malcolm Ree, (210) 536-6787

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A study was conducted to investig[illegible] the nexus of cognitive and psychomotor tests which might be used for personnel selection and assessment. These [illegible] domains are frequently seen as independent. A multiple aptitude cognitive test battery and a psychomotor te[illegible]attery were administered to 354 United States Air Force recruits. The average multiple correlation of the cogni[illegible] tests and each psychomotor score as a criterion was 0.34, corrected for range restriction. Confirmatory factor analyses disc[illegible]d general cognitive and general psychomotor factors, three lower-order psychomotors, and two lower-order cognitive fac[illegible]. The general cognitive factor accounted for 39% of the variance and the general psychomotor factor accounted for [illegible]% of the variance. Residualized, the lower-order factors accounted for between 10% and 3% of the variance. The average [illegible] saturations (loadings) of the cognitive and psychomotor tests were 0.82 and 0.34, respectively. An implication for personnel selection is that the incremental validity of psychomotor tracking tests beyond the validity of cognitive tests will be small due to the commonality of measurement. A further implication of findings is the need to study the validity of the general and specific psychomotor factors.

14. SUBJECT TERMS

Confirmatory factor analysis
General cognitive ability
Psychomotor ability

15. NUMBER OF PAGES 19

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

This effort was performed under work unit 77192501 in support of research and development of selection and classification procedures for U.S. Air Force aircrew personnel. This paper was previously published in 1994 in *The International Journal of Selection and Assessment, 2*, 209-216.

CORRELATION OF GENERAL COGNITIVE ABILITY AND PSYCHOMOTOR TRACKING TESTS

Malcolm James Ree
Thomas R. Carretta

Summary

A study was conducted to investigate the nexus of cognitive and psychomotor tests as might be used for personnel selection and assessment. These two domains are frequently seen as independent. A multiple aptitude cognitive test battery and a psychomotor test battery were administered to 354 United States Air Force recruits. The average multiple correlation of the cognitive tests and each psychomotor score as a criterion was 0.34, corrected for range restriction. Confirmatory factor analyses disclosed general cognitive and general psychomotor factors, three lower-order psychomotor factors, and two lower-order cognitive factors. The general cognitive factor accounted for 39% of the variance and the general psychomotor factor accounted for 29% of the variance. Residualized, the lower-order factors accounted for between 10% and 3% of the variance. The average *g* saturations (loadings) of the cognitive and psychomotor tests were 0.82 and 0.34, respectively. An implication for personnel selection is that the incremental validity of psychomotor tracking tests beyond the validity of cognitive tests will be small due to the commonality of measurement. A further implication of findings is the need to study the validity of the general and specific psychomotor factors.

Introduction

Most psychologists view cognitive and psychomotor abilities as distinct and independent categories (Fleishman and Quaintance, 1984, p. 162). Multiple-aptitude batteries, cognitive measures, are frequently factored and among the factors reported are verbal, quantitative, spatial, perceptual speed, technical information, and higher-order psychometric *g* (Jensen, 1980; Kass, Mitchell, Grafton, & Wing, 1983; Ree & Earles, 1991a; Skinner and Ree, 1987). Psychomotor batteries have yielded factors including control precision, multi-limb coordination, reaction time, and rate control (Fleishman, 1953, 1964, 1966, 1972), but no higher-order factor like psychometric *g*. Fleishman (1964) and Cronbach (1970) have both stated that, unlike cognitive tests, there is no higher-order general psychomotor factor.

The lack of similarity of factor names from the cognitive and psychomotor domains may represent different theoretical and taxonomic perspectives. It is possible that the same constructs are being referred to by different names. Further, it is possible that a higher-order psychomotor factor exists. The current study was conducted with the goal of investigating these issues in the context of measurement of current status as might be assessed during personnel selection. Another way of investigating cognitive and psychomotor constructs would be in their acquisition over extended time (Ackerman, 1988), such as during training or maturation. Fleishman and others (Fleishman & Hempel, 1954, 1955; Reynolds, 1952) have reported positive correlations

between cognitive and psychomotor tests which appear to vary during specific stages of psychomotor skill acquisition. However, it is unlikely that selecting agencies would be willing to expend the numerous hours necessary to examine skill acquisition. For this reason, the present study was done to estimate the correlation between cognitive abilities and current status on psychomotor variables rather than rate, variance, and maximum level of psychomotor skills acquired.

Psychometric *g* has frequently been shown to be important in the prediction of job-related criteria. For example, Hunter and Hunter (1984) showed that *g* was valid for all jobs in a large-scale meta-analysis of the United States Employment Service validity data base. Thorndike (1986) showed the predictiveness of *g* for several criteria including high school grades, Army technical training performance, job performance, and pass-fail in military pilot training. More recently, Ree and Earles (1991a) showed the predictiveness of *g* for technical training in the military and confirmed Hunter and Hunter's (1984) finding of little variation in the relationship of *g* to criteria across jobs. Ree and Earles (1992), in reviewing several studies, also found that *g* was predictive of work sample performance, supervisory ratings, oral exams of technical knowledge and procedures, graduation from pilot training, ability to perform flying maneuvers, and ability to correctly use celestial navigation equipment.

Hunter and Hunter (1984) in a reanalysis of Ghiselli's (1973) work on the mean validity of predictors, demonstrated that psychomotor scores were predictive across a wide variety of job families. They also demonstrated incremental validity of psychomotor scores when combined in a multiple regression with measures of *g*. Likewise, McHenry, Hough, Toquam, Hanson, and Ashworth (1990) found incremental validity beyond *g* measures for psychomotor tests across numerous criteria ranging from job proficiency to personal discipline and fitness. Like Hunter and Hunter, they also noted that the validity increment was small. Psychomotor scores have been found to be internationally useful in pilot selection (Carretta, 1989, 1990, 1992b; Gibb & Dolgin 1989; Louw, 1987; Pascual, 1975). Additionally, they have been evaluated as selection instruments for automobile mechanics (Porret and Frischknecht, 1975), soap packers (Shanthamani, 1979), sewing machine operators (Inskeep, 1971), and police and fire fighters (Johnson, 1984).

Hunter (1980) demonstrated some commonality between paper-and-pencil test scores on the General Aptitude Test Battery (GATB) and unrefined manual GATB psychomotor scores. However, the GATB psychomotor tests appear substantially different from those being developed today (Carretta, 1989) for use in personnel selection. These new psychomotor tests allow precise computer measurement and require manipulation of control sticks rather than the simple manual dexterity required by the GATB psychomotor tests.

Although studies of the factor structure of cognitive tests (Ree & Carretta, 1994; Ree, Mullins, Mathews, & Massey, 1982; or psychomotor tests (Fleishman, 1964, 1966) were available, no studies could be found which simultaneously investigated the factor structure of several cognitive and psychomotor tests. A joint factor analysis of cognitive and psychomotor

tests might disclose if they were factorially similar. This factorial similarity could be the reason that psychomotor tests have demonstrated only small incremental validity in personnel selection. Additionally, no empirical investigations of a general psychomotor factor could be found. The purposes of this study were: to investigate the joint factor structure of general cognitive ability measures and a typical group of computer-based psychomotor tracking tests which have been used for personnel selection; to estimate the *g* saturation of the cognitive and psychomotor tests; and to determine if a general psychomotor factor existed.

Method

Subjects

The subjects were a random sample of 354 United States Air Force recruits with a median age of 21 years and were mostly White (78%), male (86%), and high school graduate or better (99%). All subjects were selected in large part on the basis of their aptitude scores and educational achievement.

Measures

The Armed Services Vocational Aptitude Battery (ASVAB) is a 10 test multiple-aptitude measure used by the American military. It measures psychometric *g* (Ree & Earles, 1991a) and factors found to be valid for predicting a variety of criteria (Earles & Ree, 1992; Ree & Earles, 1991a, 1992). The tests include: General Science (GS), Arithmetic Reasoning (AR), Word Knowledge (WK), Paragraph Comprehension (PC), Numerical Operations (NO), Coding Speed (CS), Auto and Shop Information (AS), Mathematics Knowledge (MK), Mechanical Comprehension (MC), and Electronic Information (EI). The NO and CS tests were speeded, all the others are power. Psychometric *g* was computed as a hierarchical factor (Ree & Earles, 1991b). The AR, MK, WK, and PC tests were used to estimate psychometric *g*. Because they are based on specific content which is not part of all curricula, the other six tests might yield inadequate or biased estimates of *g* and were therefore omitted from the study.

The psychomotor tests were from the Basic Attributes Test (BAT) battery which has been validated for selection of candidates for United States Air Force pilot training (Carretta, 1989, 1990, 1992a). These tests or minor variants are frequently used by North Atlantic Treaty Organization (NATO) countries for pilot selection. The BAT was computer administered with a special alpha-numeric keypad, a monochrome monitor, and two control (joy) sticks. The first psychomotor test was a pursuit tracking task called Two-Hand Coordination, an example of Fleishman's multi-limb coordination (Fleishman & Quaintance, 1984). In this test, the subject used right- and left-hand control sticks to keep a circle on a representation of an airplane as it moved in an ellipse on the computer monitor. The two scores computed were horizontal tracking distance error (THH) and vertical tracking distance error (THV). Complex Coordination, an example of control precision and multi-limb coordination (Fleishman & Quaintance, (1984) was

the second psychomotor test. Using the right-hand control stick, this compensatory tracking task required the subject to keep a 1" cross centered on a dotted-line cross that bisected the monitor horizontally and vertically. Simultaneously, using the left-hand control stick, the subject has to keep a 1" vertical bar horizontally centered at the base of the monitor display. The 1" cross and the vertical bar were forced away from center by a random function. The three scores for this test were: horizontal tracking distance error (CCH) and vertical tracking distance error (CCV) for the 1" cross, and tracking distance error (CCR) for the 1" vertical bar. The third psychomotor test, Time Sharing, was identified with Fleishman and Quaintance's (1984) psychomotor factors of reaction time and rate control. In the first 10 minutes, the subject was required to keep randomly-moving cross-hairs on an airplane target using the right-hand control stick. In the next 6 minutes, the subject had to repeat the tracking task and had to cancel digits which appeared at random intervals and positions on the monitor. Cancellation was timed and consisted of pressing the corresponding digit on the numeric keypad. Tracking task difficulty was computer adjusted. Smaller tracking errors caused the stick sensitivity to increase and larger tracking errors caused it to decrease. The three scores on this test were: tracking difficulty on the task without digit cancellation (TSS), digit cancellation reaction time (TSR), and tracking difficulty during digit cancellation (TSD). Correlations involving error and response time scores were reflected so that good performances were always positively correlated.

Electro-mechanical versions of these psychomotor tests were administered during the second World War and are reported by Thorndike and Hagen (1959). McGrevy and Valentine (1974) designed the current computer-based tests to require the same subject responses as the World War II versions of the tests. A detailed description of the BAT was provided by Carretta (1987).

Procedures

The cognitive tests were administered as part of the operational enlistment qualification procedures and the BAT was administered on the 11th day of basic military training. The subjects were told that the BAT scores were being collected for experimental purposes only and although given the opportunity to decline participation, none did.

Analyses included descriptive statistics, correlations, regressions, and factor analyses. All correlations were tested at $\underline{p} < 0.01$ Type I error rate.

Because the sample was selected on the basis of minimum scores on the cognitive ability tests, the range of variance on all variables has been curtailed. In the early part of the century, Pearson (1903) observed that correlations computed in such range-restricted samples were substantially downwardly biased estimators. He provided a simple equation to correct the estimates when only two variables (i.e., the predictor and the criterion) were involved. Thorndike (1949) popularized the procedure and provided a dramatic example of how misinterpretation could occur if uncorrected correlations were used rather than corrected correlations. Thorndike showed range restriction caused by selection downwardly biased the correlational estimates

including turning some correlations negative when they were positive in the unrestricted sample. Ree, Carretta, Earles, and Albert (1994) have demonstrated that the correction procedures can return the appropriate sign to correlations.

Lawley (1943) provided the general solution for the correction of multivariate range restriction. For example, if minimum scores were required on four tests for selection into a job, multivariate restriction occurred. Lawley's procedure is applicable in such cases and corrects for multivariate restriction in range.

Linn, Harnisch, and Dunbar (1981) have demonstrated that corrected correlations are better estimates of the population value than are uncorrected range-restricted correlations. Further, they demonstrated that the corrected correlations are conservative estimates of the population correlations. Clearly, the corrected correlations are better statistical estimates of the population parameters and should be preferred to the biased uncorrected correlations.

The matrix of correlations of cognitive and psychomotor tests was corrected for range restriction by the multivariate method (Birnbaum, Paulson, & Andrews, 1950; Lawley, 1943). Each psychomotor test score was predicted by the set of four cognitive tests as a measure of commonality with cognitive ability. The four cognitive tests display multicolinearity that would increase the sampling variance of the regression weights. However, multicolinearity creates no problem for the estimation of the multiple correlation coefficient.

The first higher-order common factor from a confirmatory factor analysis was extracted from the corrected correlation matrix to estimate *g* loadings (Jensen, 1980; Ree & Earles, 1991a) of all the tests. To determine if the first factor was still *g*, as in previous studies (Ree & Carretta, 1994), the loadings for the cognitive tests were estimated without the psychomotor tests present and compared to the loading found when the psychomotor tests were present.

Multiple confirmatory factor analyses were conducted on the range-restriction corrected correlation matrix to find a parsimonious model which fitted the data. The first was a *g*-only model specifying that one factor represented the data. The second model posited *g* from the four cognitive tests and a psychomotor factor from the eight psychomotor scores. These factors were allowed to be correlated. Model three specified *g* coming from the four cognitive tests and three psychomotor factors each representing scores from a particular psychomotor test. Model four extracted a verbal factor from PC and WK, a quantitative factor from MK and AR, and three psychomotor factors, one for each psychomotor test (i.e., Two-Hand Coordination, Complex Coordination, and Time Sharing). Model five was model four with a hierarchical *g* and a hierarchical psychomotor factor added. Although starting with correlated factors, model five was completely residualized (Schmit & Leiman, 1957) so that the effects of *g* and the hierarchical psychomotor factor were removed from the lower-order factors.

Several fit statistics including the goodness-of-fit χ^2, Bentler-Bonett non-normed fit statistic (Bentler, 1989), Tucker-Lewis fit index (Tucker & Lewis, 1973), and the standardized

residuals were evaluated for each model to determine which was most appropriate. Marsh, Balla, and McDonald (1988) have shown that the Bentler-Bonnet index may be susceptible to sample size effects. They recommend evaluation of the Tucker-Lewis TLI incremental fit index to determine the most appropriate factor structure. Bentler (1990) has developed the Comparative Fit Index (CFI) based on the TLI and has shown that it is less dependent on sample size and has a smaller sampling variance than the TLI. The CFI was used to evaluate models for goodness-of-fit.

Direct comparison between 'nested' models was also accomplished by evaluating the difference between goodness-of-fit χ^2 statistics and only accepting a model as better if a significant χ^2 difference was found. This comparison could only be made for models one versus five, and four versus five.

Results

Computation of descriptive statistics of the cognitive tests showed that the sample was range restricted. On average, cognitive test scores were about one-half standard deviation above the normative mean and average variances, one-fourth the normative variances. Table 1 shows the correlations among the cognitive and psychomotor tests, both as observed and as corrected for range restriction.

Table 1. Correlation of cognitive and psychomotor tracking tests

	AR	WK	PC	MK	THH	THV	CCH	CCV	CCR	TSR	TSS	TSD
AR	1.00	0.32	0.31	0.59	0.19	0.20	0.15	0.22	0.19	0.23	0.18	0.19
WK	0.71	1.00	0.53	0.26	0.14	0.14	0.06	0.11	0.07	0.12	0.08	0.04
PC	0.67	0.80	1.00	0.27	0.07	0.06	-0.01	0.00	0.00	0.14	0.05	0.01
MK	0.83	0.67	0.64	1.00	0.17	0.17	0.13	0.16	0.17	0.17	0.22	0.19
THH	0.36	0.31	0.25	0.31	1.00	0.92	0.38	0.43	0.33	0.23	0.51	0.48
THV	0.34	0.29	0.22	0.30	0.93	1.00	0.41	0.44	0.38	0.21	0.50	0.48
CCH	0.26	0.17	0.10	0.22	0.42	0.44	1.00	0.64	0.57	0.31	0.51	0.50
CCV	0.37	0.28	0.17	0.31	0.49	0.49	0.67	1.00	0.57	0.25	0.48	0.48
CCR	0.33	0.24	0.17	0.29	0.39	0.43	0.60	0.61	1.00	0.27	0.38	0.38
TSR	0.46	0.40	0.41	0.41	0.33	0.30	0.35	0.34	0.35	1.00	0.37	0.44
TSS	0.27	0.17	0.14	0.27	0.54	0.53	0.53	0.51	0.42	0.41	1.00	0.85
TSD	0.24	0.12	0.09	0.22	0.50	0.50	0.52	0.51	0.41	0.46	0.86	1.00

Note: Entries above the diagonal are observed data, those below have been corrected for range restriction. Among the uncorrected correlations any value exceeding 0.13 is statistically significant at $p < 0.01$;

however, because range restriction drastically reduces statistical power, the statistical test is of little value in rejecting the null hypothesis that $\rho = 0$.

Contrary to expectations from the literature, significant correlations were observed between psychomotor and cognitive scores. The average corrected for range-restriction correlation between cognitive test scores and psychomotor test scores was 0.26. The Arithmetic Reasoning test had the highest average correlation with the psychomotor scores at 0.33; while Paragraph Comprehension had the lowest correlation at 0.19. Among the psychomotor scores, the Time Sharing test provided both the most and least correlated with the cognitive tests. TSR, Time Sharing response time, was the most correlated at 0.42 and TSD, Time Sharing difficulty with digit cancellation, the least at 0.17.

Table 2 presents the results of the commonality analyses of the psychomotor tests. Commonality was estimated by regressing each psychomotor test score on the paper-and-pencil cognitive tests. All correlations were statistically significant. The average multiple correlation of psychomotor scores predicted by paper-and-pencil test scores was 0.22. Corrected for range restriction, the average correlation (R_c) was 0.34.

Table 2. Multiple correlations of psychomotor tests with cognitive tests

Test	R	R_c
THH	0.21	0.37
THV	0.22	0.35
CCH	0.17	0.28
CCV	0.24	0.40
CCR	0.21	0.34
TSR	0.24	0.48
TSS	0.22	0.29
TSD	0.21	0.26

Note. R is the multiple correlation between the psychomotor test and the cognitive tests and R_c is R corrected for range restriction. All observed correlation R were significant at p <0.01 Type I error rate

The five confirmatory factor analytic models were estimated using maximum likelihood procedures (Bentler, 1989). The *g*-only model showed a poor fit to the data with a CFI of 0.459 (χ^2) = 1772, df = 53, $p < .01$). The fit of the two-factor model two, *g* from the paper-and-pencil tests and one factor from the psychomotor tests, was also poor at 0.702 ($\chi^2 = 999$, df = 51, $p <$ 0.01). The four-factor model extracting *g* from only the paper-and-pencil tests, and three

psychomotor factors gave a better CFI of 0.933 ($\chi^2 = 260$, df = 46, $p < 0.01$). A five-factor model including verbal and math and three psychomotor factors showed a CFI of 0.970 ($\chi^2 = 136$, df = 40, $p < 0.01$). Adding a hierarchical *g* and hierarchical psychomotor factor to the five-factor model increased the fit to yield a CFI of 0.993 ($\chi^2 = 57$, df = 34, $p < 0.01$). The χ^2 difference between models one and five was 1715 (df = 19, $p < 0.01$), and the χ^2 difference between models four and five was 79 (df = 6, $p < 0.01$). Clearly, model five did not add parameters needlessly. Model five yielded the best fit by all indexes.

The results revealed a substantial first higher-order factor with all the cognitive tests loaded positively, suggesting that it was an estimate of psychometric *g*. Additionally, each of the psychomotor tests had positive loadings on this factor. To determine if the first higher-order factor were still a measure of *g* with the inclusion of the psychomotor tests, the first-order factor loadings of the cognitive tests were estimated without the presence of the psychomotor tests. The two sets of loadings were almost identical with no differences greater than 0.04, suggesting that the factor being measured was the same in each case.

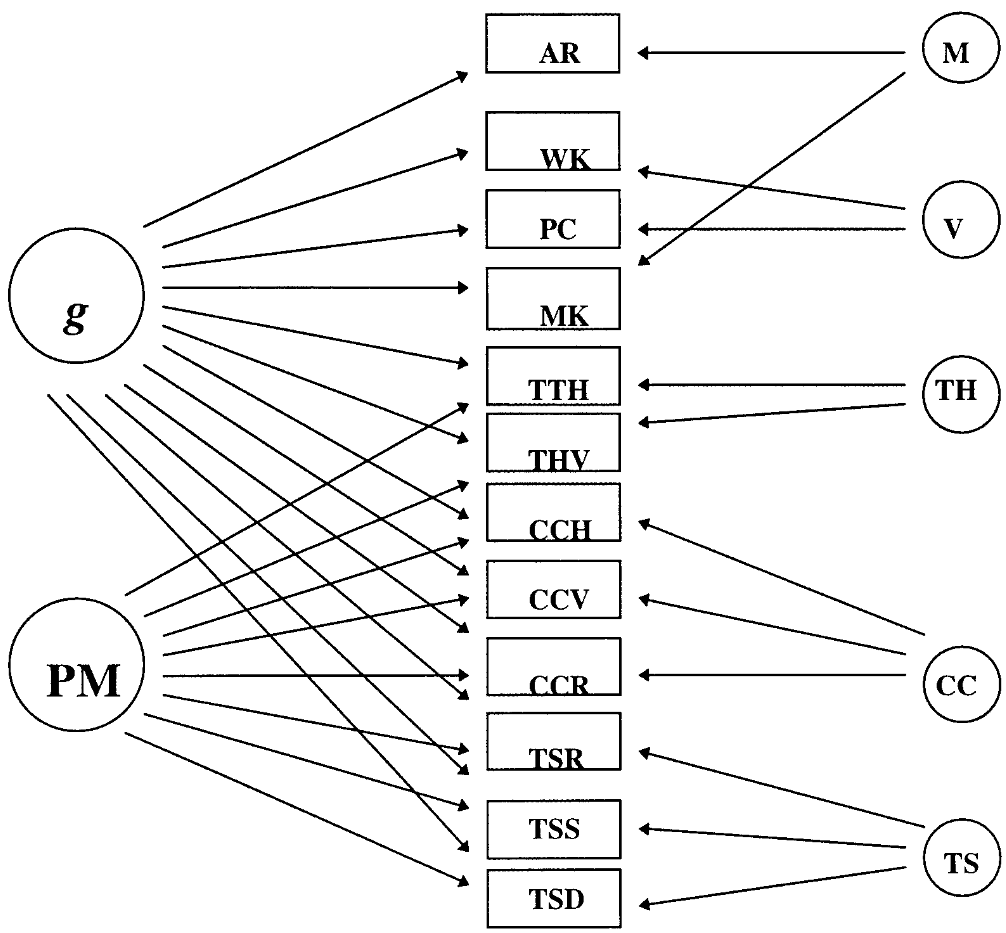

Figure 1. Structural model of test performance from confirmatory factor analysis.

Figure 1 shows the factor structure for the cognitive and psychomotor tests. The factors were interpreted as psychometric *g*, a higher-order general psychomotor factor (PM), a two-hand coordination factor (TH), a complex coordination factor (CC), a time sharing factor (TS), a verbal factor (V), and a mathematical factor (M). The proportions of the variance attributable to the higher-order factors were 39% for *g* and 29% for general psychomotor. Among the residualized lower-order factors, two-hand coordination accounted for 10%; complex coordination, 7%; time sharing, 7%; verbal, 5%; and mathematical, 3%.

The *g* loadings for the psychomotor tests were lower than those for the cognitive tests. The ratio of the average *g* saturation of the cognitive (0.82) and psychomotor tests (0.34) was a little more than two to one (Table 3).

Arithmetic Reasoning showed the highest *g* loading, 0.88, among the paper-and-pencil cognitive tests, and Time Sharing reaction time (TSR) showed the highest *g* loading, 0.51, among the psychomotor tests. The range of the *g* loadings among the paper-and-pencil tests was modest at 0.12, while the range for the psychomotor tests was more than double at 0.29. The lowest *g* loading among the cognitive tests was greater than the highest *g* load among the psychomotor tests.

Table 3. Loadings of the confirmatory factor analysis

	Factor						
	I	II	III	IV	V	VI	VII
Test	*g*	PM	TH	CC	TS	V	M
AR	0.88						0.45
WK	0.80					0.29	
PC	0.76					0.65	
MK	0.82						0.21
THH	0.38	0.57	0.62				
THV	0.36	0.58	0.72				
CCH	0.25	0.62		0.47			
CCV	0.36	0.58		0.45			
CCR	0.34	0.46		0.48			
TSR	0.51	0.26			0.23		
TSS	0.27	0.73			0.40		
TSD	0.22	0.73			0.64		
Percent	39	29	10	7	7	5	3

Note. The factors are: *g* is psychometric *g*, PM is higher-order psychomotor, TH is two-hand psychomotor, CC is complex coordination psychomotor, and TS is time sharing psychomotor, V is verbal, and M is mathematical. *g* and PM are higher-order factors and the others are residualized lower-order factors and all are orthogonal after residualization. Percentages are given as rounded integers.

Discussion

The finding that psychomotor and cognitive tests had correlated scores was contrary to the expectancy derived from the literature, especially the work of Fleishman and others (Fleishman, 1953, 1964, 1966, 1972; Fleishman & Hempel, 1954, 1955; Fleishman & Quaintance 1984; Reynolds, 1952). It is easy to understand how cognitive tests and psychomotor tests could be seen to be relatively independent by inspection of the uncorrected correlations in Table 1. However, the corrected correlations showed that the two types of measures were not totally independent. The correlations between cognitive and psychomotor tests may be due to the requirement to reason (the foundation of *g*) while taking the tests.

The current study was not concerned with the acquisition of either cognitive or psychomotor skills, but rather with the relationships of scores collected at some point in time. This would be consistent with how such scores could be used in personnel selection. It is possible that the correlations among cognitive and psychomotor tests could change during or after acquisition of skill levels (Ackerman, 1988; Fleishman & Hempel, 1954, 1955; Reynolds, 1952).

The finding that psychomotor tests were *g*-loaded was generally unexpected. Given the importance of *g*, (Hunter & Hunter, 1984; Ree & Earles, 1991a; Thorndike, 1986) the implication of this finding is that the incremental predictive validity of psychomotor tests (Carretta & Ree, 1994) will not be as great for many jobs as suggested by the psychomotor literature. This is because the *g* component of psychomotor scores is not independent of the *g* component of paper-and-pencil tests. The amount of incremental validity provided by the non-*g* portions of psychomotor test scores will be proportional to the amount of non-*g* psychomotor variance in the criterion.

Contrary to the assertions of Fleishman (1964) and Cronbach (1970), the confirmatory factor analysis showed a general (higher-order) psychomotor factor. As in cognitive tests, this factor might be the major source of validity, as might the three lower-order psychomotor factors or some combination of higher-and lower-order factors. The validity of these factors could be important in the development of alternative psychomotor tests forms for use in personnel selection systems. If the main predictive portion of the psychomotor tests were the higher-order factor (PM) as in cognitive tests, then most aggregations of psychomotor tests like those evaluated in the current study, could be expected to provide a measure of general psychomotor ability. This would facilitate building alternative or replacement psychomotor test forms. If, however, the validity were a consequence of only the specific psychomotor factors then the development of alternate forms would be more difficult as new measures of the specific lower-order factors were sought.

Other psychomotor factors such as those described by Fleishman and Quaintance (1984) should be developed into tests and evaluated by the methods used here. Further, the validity of general and specific psychomotor factors must be studied in the same manner (Ree & Earles,

1991a) as the validity of general and specific cognitive factors to understand their role in prediction.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology; General, 118*, 288-318.

Bentler, P. M. (1989). EQS *Structural Equations Program Manual.* Los Angeles: CA, BMDP Statistical Software.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.

Birnbaum, Z. W., Paulson, E., & Andrews, F. C. (1950). On the effect of selection performed on some coordinates of a multi-dimensional population. *Psychometrika, 15*, 191-204.

Carretta, T. R. (1987). *Basic Attributes Test (BAT) system: The development of an automated test battery for pilot selection* (AFHRL-TR-87-9). Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Carretta, T. R. (1989). USAF pilot selection and classification systems. *Aviation Space and Environmental Medicine, 60*, 46-49.

Carretta, T. R. (1990). Cross-validation of experimental USAF pilot training performance models. *Military Psychology, 2*, 257-264.

Carretta, T. R. (1992a). Short-term retest reliability of an experimental US Air Force pilot candidate selection test battery. The *International Journal of Aviation Psychology, 2* 161-173.

Carretta, T. R. (1992b). Understanding the relations between selection factors and pilot training performance: Does the criterion make a difference? *The International Journal of Aviation Psychology, 2*, 95-105.

Carretta, T. R., & Ree, M. J. (1994). Pilot candidate selection method (PCSM): Sources of validity. *The International Journal of Aviation Psychology, 4*, 103-117.

Cronbach, L. J. (1970). *Essentials of Psychological Testing* (3rd ed.). New York: Harper and Row.

Earles, J. A., & Ree, M. J. (1992). The predictive ability of the ASVAB for training grades. *Educational and Psychological Measurement, 52*, 721-725.

Fleishman, E. A. (1953). Testing for psychomotor abilities by means of apparatus tests. *Psychological Bulletin, 50*, 241-262.

Fleishman, E. A. (1964). *The Structure and Measurement of Physical Fitness*. Englewood Cliffs, NJ: Prentice-Hall.

Fleishman, E. A. (1966). Human abilities and the acquisition of skill. In E. A. Bilodeau (ed.), *Acquisition of Skill*. New York: Academic Press.

Fleishman, E. A. (1972). On the relation between abilities, learning, and human performance. *American Psychologist, 27*, 1017-1032.

Fleishman, E. A., & Hempel, W. E., Jr. (1954). Changes factor structure of a complex psychomotor test as a function of practice. *Psychometrika, 19*, 239-252.

Fleishman, E. A., & Hempel, W. E., Jr. (1955). The relation between abilities and improvement with practice in a visual discrimination reaction task. *Journal of Experimental Psychology, 49* 301-312.

Fleishman, E. A. & Quaintance, M. K. (1984). *Taxonomies of Human Performance: The Description of Human Tasks*. Orlando, FL: Academic Press.

Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology, 26*, 461-477.

Gibb, G. D., & Dolgin, D. L. (1989). Predicting military flight training success by a compensatory tracking task. *Military Psychology, 1*, 235-240.

Hunter, J. E. (1980). *The Dimensionality of the General Aptitude Test Battery (GATB) and the Dominance of General Factors over Specific Factors in the Prediction of Job Performance*. Washington, DC: US Employment Service, US Department of Labor.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*, 72-98.

Inskeep. G. C. (1971). The use of psychomotor tests to select sewing machine operators: Some negative findings. *Personnel Psychology, 24*, 707-714.

Jensen, A. R. (1980). *Bias in Mental Testing*. New York: The Free Press.

Johnson, E. E. (1984). Problems in assessing police and fire fighter candidates. *Journal of Police Science and Administration, 12*, 404-406.

Kass, R. A., Mitchell, K. J., Grafton, F. C., & Wing, H. (1983) Factorial validity of the Armed Services Vocational Aptitude Battery (ASVAB), Forms 8, 9, and 10. *Educational and Psychological Measurement, 43*, 1077-1087.

Lawley, D. N. (1943). A note on Karl Pearson's selection formulas. *Proceedings of the Royal Society of Edinburgh.* Section A, *62*, Part I, 28-30.

Linn, R. L., Harnisch, D. L., & Dunbar, S. (1981). Corrections for range restriction: An empirical investigation of conditions resulting in conservative corrections. *Journal of Applied Psychology, 66*, 655-663.

Louw, J. (1987). World War II, industry, and the professionalization of South African psychology. *South African Journal of Psychology, 17*, 33-39.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*, 391-410.

McGrevy, D. F., & Valentine, L. D., Jr. (1974). *Validation of two aircrew psychomotor tests (AFHRL-TR-74-4).* Lackland AFB, TX: Personnel Research Division, Air Force Human Resources Laboratory.

McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology, 43*, 335-354.

Pascual, M. (1975). Structure and dimension of aptitude for flying. *Revista de Psicologia-General y Aplicada, 30*, 133-134.

Pearson, K. (1903). Mathematical contributions to the theory of evolution-XI. On the influence of natural selection on the variability and correlation of Organs. *Royal Society Philosophical Transactions,* (Series A, 200), 1-66.

Porret, J., & Frischknect, E. (1975). The automobile mechanics' apprentices: A study of the validity of a selection procedure. *Psychologie Schweizerische Zeitschrift fur Psychologie und irhe Anwendungen, 34*, 141-159.

Ree, M. J., & Carretta, T. R. (1994). Factor analysis of ASVAB: Confirming a Vernon-like structure. *Educational and Psychological Measurement, 54*, 457-461.

Ree, M. J., Carretta, T. R., Earles, J. A., & Albert, W. (1994). Sign changes when correcting for range restriction: A note on Pearson's and Lawley's selection formulae. *Journal of Applied Psychology, 79*, 298-301.

Ree, M. J., & Earles, J. A. (1991a). Predicting training success: Not much more than *g*. *Personnel Psychology, 44*, 321-332.

Ree, M. J., & Earles, J. A. (1991b). The stability of *g* across different methods of estimation. *Intelligence, 15*, 271-278.

Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science, 1*, 86-89.

Ree, M. J., Mullins, C. J., Mathews, J. J., & Massey, R. H. (1982). *Armed Services Vocational Aptitude Battery: Item and factor analysis of Forms 8, 9, and 10* (AFHRL-TR-81-55). Brooks AFB, TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Reynolds, B. (1952). The effect of learning on the predictability of psychomotor performance. *Journal of Experimental Psychology, 44*, 189-198.

Schmid, J., & Leiman, J. (1957). The development of hierarchical factor solutions. *Psychometrika, 22*, 53-61.

Shanthamani, V. S. (1979). Industrial use of dexterity tests. *Journal of Psychological Researches, 23*, 200-205.

Skinner, J., & Ree, M. J. (1987). *Air Force Officer Qualifying Test (AFOQT): Item and factor analysis* (AFHRL-TR-86-68). Brooks AFB TX: Air Force Human Resources Laboratory, Manpower and Personnel Division.

Thorndike, R. L. (1949). *Personnel Selection*. New York: John Wiley.

Thorndike, R. L. 1986). The role of general ability in prediction. *Journal of Vocational Behavior, 29*, 332-339.

Thorndike, R. L., & Hagen, E. (1959). *Ten Thousand Careers*. New York: Wiley.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1-10.